

Lower Bounds for Ground States of Condensed Matter Systems

Tillmann Baumgratz and Martin B. Plenio

Institut für Theoretische Physik, Albert-Einstein-Allee 11, Universität Ulm, D-89069 Ulm, Germany

(Dated: February 28, 2012)

Standard variational methods tend to obtain upper bounds on the ground state energy of quantum many-body systems. Here we study a complementary method that determines lower bounds on the ground state energy in a systematic fashion, scales polynomially in the system size and gives direct access to correlation functions. This is achieved by relaxing the positivity constraint on the density matrix and replacing it by positivity constraints on moment matrices, thus yielding a semi-definite programme. Further, the number of free parameters in the optimization problem can be reduced dramatically under the assumption of translational invariance. A novel numerical approach, principally a combination of a projected gradient algorithm with Dykstra's algorithm, for solving the optimization problem in a memory-efficient manner is presented and a proof of convergence for this iterative method is given. Numerical experiments that determine lower bounds on the ground state energies for the Ising and Heisenberg Hamiltonians confirm that the approach can be applied to large systems, especially under the assumption of translational invariance.

I. BACKGROUND AND MOTIVATION

The determination of the ground state properties of strongly interacting quantum systems is of considerable interest. As only a very few models are exactly solvable, numerical approximation methods are of significant importance. A key challenge in this respect is the ‘curse of dimensionality’, namely the exponential growth of the Hilbert space dimension with the number of sites. Ingenious methods such as the density matrix renormalization group (DMRG) approach [1, 2] can overcome this problem as they amount to formulating a variational problem over a well-chosen class of quantum states that can be parametrized efficiently, i.e. with a polynomial number of parameters [3]. As the variation is over a subclass of all possible quantum states, variational methods of this type provide upper bounds on the ground state energy. Assessing the quality of the so obtained upper bounds on the ground state energy is then difficult. Hence there is considerable interest in the development of efficient methods for the determination of lower bounds on the ground state energy. One straightforward approach in this direction consists in the determination of Anderson bounds [4], i.e. the decomposition of a Hamiltonian $H = \sum_k H_k$ and the subsequent determination of the ground state energy of each H_k whose sum then yields a lower bound on the ground state energy of H . While useful, this approach does not scale well with the system size as it remains exponential in the size of the support of the H_k . The application of DMRG to the H_k may overcome the scaling issue but amounts to finding an ‘upper bound to a lower bound’ and does not resolve the principal issue with DMRG.

As early as 1940, however, it was suggested to avoid the use of wave functions altogether and rather consider reduced density matrices together with conditions that these arise from a valid quantum state [5] – the so-called N-representability problem [6, 7]. Varying over the reduced density matrices that are compatible with a valid physical state then allows for the determination of the ground state energy. The exact ground state energy can be obtained if all conditions for the N-representability problem are known. If only partial conditions are specified, then one minimizes over too large a set of reduced density matrices and one obtains a lower bound on the ground state energy. The analytical treatment of this problem is highly challenging and at the time numerical solutions were hard to come by. Further progress was made by Mazziotti [8, 9], who pointed out that the problem can be formulated as a semi-definite programme and use can be made of the guaranteed and certified convergence of the primal-dual interior-point algorithm for semi-definite programming [10, 11]. A general mathematical framework for this type of optimization problems and applications was given in [12]. Unfortunately, the primal-dual interior-point algorithm and related methods do come with a considerable overhead in computation time and, crucially, memory that is making it hard to scale to very large systems. Lately, new algorithms to address the optimization problems of finding lower bounds on ground state energies have been developed [13, 14] and considerable improvements towards the interior-point algorithms have been demonstrated.

In this work, we are exploring further the idea of relaxations of the ground state energy finding problem that allow us to determine lower bounds on the ground state energy in an efficient manner. To this end we formulate the problem, introduce a number of simplifications including the explicit treatment of translation invariance and symmetries and point out additional potential routes for numerical simplifications. In addition, we present a novel algorithm, principally a combination of a projected gradient algorithm with Dykstra’s algorithm [15–17], for solving the resulting optimization problem in a memory-efficient manner and prove its convergence. Then, in order to demonstrate the capabilities of this approach, we present numerical experiments that determine lower bounds on the ground state energies for two one-dimensional (1D) lattice models, namely the Ising and Heisenberg Hamiltonian, which confirm that the approach can be applied for large systems especially when implementing translational invariance. Note that although we restrict our numerical analysis to 1D lattice models it has been demonstrated that the variational determination of ground state energies considering only constraints on reduced density matrices applies to higher spatial dimensions [18] and to non-periodic systems such as molecules in quantum chemistry [13] as well. We complete this work with a discussion and outlook.

II. BASICS

We begin by explaining the general ideas underlying the method for the determination of lower bounds on ground state energies following [8] and then specialize to the case of fermionic systems, for which later sections will provide numerical examples. Analogous ideas can also be implemented for bosonic systems or general spin-systems.

A. Generalities

The determination of the ground state energy E_0 of a many-body system composed of N particles and described by the Hamilton operator H can be formulated as the minimization problem

$$E_0 = \min \{ \text{tr} [H\rho] \mid \rho \geq 0, \text{tr} [\rho] = 1 \}. \quad (1)$$

This approach is pushed to its limits rather rapidly as it uses the positivity constraint on ρ , which grows dramatically with the system size, e.g. for particles with spin-1/2 or spinless fermions ρ is a $2^N \times 2^N$ matrix. Hence, this problem cannot be solved efficiently due to the exponentially increasing parameter space. One technique to deal with this difficulty is to parametrize a specific family of candidates $|\phi(\vec{\alpha})\rangle$ for the ground state by a set of parameters $\vec{\alpha}$. One then minimizes the expectation value of the Hamiltonian with respect to this family of states. As this family of states does not encompass the entire state space, this method will find a state $|\phi_0\rangle$ which is optimal in $\{|\phi(\vec{\alpha})\rangle\}_{\vec{\alpha}}$ and which yields an upper bound on the true ground state energy. This is the route taken by a variety of variational methods including, for example, DMRG [1–3].

To overcome the scaling issues and to provide lower rather than upper bounds on the ground state energy, we need to relax the positivity constraint $\rho \geq 0$, replacing it by a weaker set of constraints which scale only polynomially in the system's size [8, 12]. This enlarges the class of states over which one varies the energy and hence yields lower bounds on the ground state energy. To this end we use the fact that

$$\rho \geq 0 \Rightarrow \text{tr} [\mathcal{O}^\dagger \mathcal{O} \rho] \geq 0 \quad (2)$$

for all operators \mathcal{O} . Demanding only

$$\langle \mathcal{O}^\dagger \mathcal{O} \rangle = \text{tr} [\mathcal{O}^\dagger \mathcal{O} \rho] \geq 0 \quad (3)$$

for some operator \mathcal{O} does not, in general, imply $\rho \geq 0$. For some subset of operators $\{O_k\}_k$, we can then define $\mathcal{O} = \sum_k \alpha_k O_k$ and the requirement that $\langle \mathcal{O}^\dagger \mathcal{O} \rangle \geq 0$ for all choices of α_k is then equivalent to the positivity of the matrix X with components $X_{kl} = \langle O_k^\dagger O_l \rangle$, imposing a particular N-representability condition. Hence, replacing in the original optimization problem eq. (1) the condition $\rho \geq 0$ with the weaker condition $X \geq 0$ and omitting the normalization $\text{tr}[\rho] = 1$, we find that

$$E_0 \geq \min \{ \text{tr} [H\rho] \mid X \geq 0 \}. \quad (4)$$

The present formulation still contains the density operator ρ explicitly in the computation of $\text{tr} [H\rho]$. Judicious choice of the set $\{O_k\}_k$ alleviates this shortcoming. Indeed, writing $H = \sum_{kl} h_{kl} H_{kl}$, where $h_{kl} \in \mathbb{C}$ are the coefficients of the operators H_{kl} in the Hamiltonian H , one ensures that the set $\{O_k\}_k$ contains all operators such that for all k, l the expectation value $\langle H_{kl} \rangle$ is in $\{\langle O_k^\dagger O_l \rangle\}_{k,l}$. Note, furthermore, that the entries of X will generally not be independent but suffer some linear equality constraints, for example due to (anti)-commutation relations between operators. Denoting by \mathcal{C} the set of all Hermitian matrices that satisfy $X \geq 0$ and any other linear constraints, the energy minimization problem can be written as [8, 9]

$$E_0 \geq \min \left\{ \sum_{kl} h_{kl} X_{kl} \mid X \in \mathcal{C} \right\}, \quad (5)$$

which is evidently a semi-definite programme [11] and efficiently solvable as long as the set $\{O_k\}_k$ contains a number of operators that scale polynomially in system size. The decomposition of H into H_{kl} , and hence the choice of the set $\{O_k\}_k$, is not unique and the optimal choice is not obvious. From a practical point of view however, the systematic choice presented in the remainder of the paper leading to the so-called p-positivity conditions [8, 9] does perhaps provide the most straightforward formulation.

B. Bosons and Fermions

For concreteness, let us now consider the class of bosonic or fermionic operators defined by

$$\mathcal{O} = \sum_k \delta_k a_k + \sum_k \epsilon_k a_k^\dagger + \sum_{kl} \alpha_{kl} a_k a_l + \sum_{kl} \beta_{kl} a_k^\dagger a_l + \sum_{kl} \gamma_{kl} a_k^\dagger a_l^\dagger, \quad (6)$$

where $\delta_k, \epsilon_k, \alpha_{kl}, \beta_{kl}, \gamma_{kl} \in \mathbb{C}$ are arbitrary complex variables. The N-representability conditions arising from this class of operators are called 1-positivity and 2-positivity conditions, which is a terminology used to reveal the powers of the creation and annihilation operators a_i^\dagger and a_i in \mathcal{O} . In general, allowing polynomials of order p in the second-quantized operators leads to restrictions called p-positivity conditions [8, 9]. For the particular choice of operators of the order of 2, requiring the positivity of $\langle \mathcal{O}^\dagger \mathcal{O} \rangle$ is equivalent to

$$X = \begin{pmatrix} T & U^\dagger & A^\dagger & B^\dagger & C^\dagger \\ U & S & D^\dagger & E^\dagger & F^\dagger \\ A & D & M & G^\dagger & H^\dagger \\ B & E & G & R & I^\dagger \\ C & F & H & I & Q \end{pmatrix} \geq 0, \quad (7)$$

where the $N \times N$ matrices comprising the second moments are defined as follows:

$$\begin{aligned} T_{k,l} &= \langle a_k^\dagger a_l \rangle, \\ S_{k,l} &= \langle a_k a_l^\dagger \rangle, \\ U_{k,l} &= \langle a_k a_l \rangle \end{aligned} \quad (8)$$

and for the third and fourth moments, we find the $N^2 \times N$ and $N^2 \times N^2$ matrices

$$\begin{aligned} A_{kl,m} &= \langle a_k^\dagger a_l^\dagger a_m \rangle, & M_{kl,mn} &= \langle a_k^\dagger a_l^\dagger a_n a_m \rangle, \\ B_{kl,m} &= \langle a_k^\dagger a_l a_m \rangle, & G_{kl,mn} &= \langle a_k^\dagger a_l a_n a_m \rangle, \\ C_{kl,m} &= \langle a_k a_l a_m \rangle, & H_{kl,mn} &= \langle a_k a_l a_n a_m \rangle, \\ D_{kl,m} &= \langle a_k^\dagger a_l^\dagger a_m^\dagger \rangle, & R_{kl,mn} &= \langle a_k^\dagger a_l a_n^\dagger a_m \rangle, \\ E_{kl,m} &= \langle a_k^\dagger a_l a_m^\dagger \rangle, & I_{kl,mn} &= \langle a_k a_l a_n^\dagger a_m \rangle, \\ F_{kl,m} &= \langle a_k a_l a_m^\dagger \rangle, & Q_{kl,mn} &= \langle a_k a_l a_n^\dagger a_m^\dagger \rangle. \end{aligned} \quad (9)$$

In addition to the positivity constraint on the $2N + 3N^2 \times 2N + 3N^2$ Hermitian matrix $X \in \mathcal{H}$, where \mathcal{H} denotes the vector space of all Hermitian matrices with appropriate dimension, there are linear constraints relating its entries. In fact, the fermionic anti-commutator relations

$$\{a_k, a_l\} = \{a_k^\dagger, a_l^\dagger\} = 0, \quad \{a_k, a_l^\dagger\} = \delta_{k,l} \quad (10)$$

or the bosonic commutator relations

$$[a_k, a_l] = [a_k^\dagger, a_l^\dagger] = 0, \quad [a_k, a_l^\dagger] = \delta_{k,l} \quad (11)$$

force some entries of X to be equal, to differ by sign or to be a linear function of others.

Now, let \mathcal{C} be the set of all positive semi-definite Hermitian matrices of the form eq. (7) and satisfying the additional requirement that the entries of X obey the constraints given by the commutation or anti-commutation relations. Further, let $E(X) : \mathcal{H} \rightarrow \mathbb{R}$ be a function expressing the expectation value of a Hamiltonian with up to 4-mode terms, i.e. $E(X) = \langle H \rangle$. Then, the original optimization problem eq. (1) can be replaced by

$$E_0 \geq \min \{E(X) \mid X \in \mathcal{C}\}, \quad (12)$$

where the relaxed optimization problem scales polynomially in the system size N . Note that it is straightforward to extend this formulation to higher moments, e.g. 3-positivity conditions. In the following, we are going to present scenarios and strategies where the number of free variables can be decreased considerably.

C. Consequences of Superselection Rules and Particle Number Conservation for Fermions

For the remainder we will consider fermionic systems. In that case, superselection rules require that the Hamiltonian commutes with the parity operator given by $P = \prod_k (1 - 2a_k^\dagger a_k)$. As a consequence, only Hamiltonians comprising terms with an even number of creation and annihilation operators are physically allowed. This in turn ensures that only expectation values of products of an even number of creation and annihilation operators are non-vanishing. Hence, we find that we can restrict X to take the form

$$X = \begin{pmatrix} T & U^\dagger & 0 & 0 & 0 \\ U & S & 0 & 0 & 0 \\ 0 & 0 & M & G^\dagger & H^\dagger \\ 0 & 0 & G & R & I^\dagger \\ 0 & 0 & H & I & Q \end{pmatrix} \quad (13)$$

and to satisfy the constraints dictated by the fermionic anti-commutator relations. See appendix B for a list of constraints on the entries for a formulation for up to fourth moments in the fermionic case.

If the Hamiltonian conserves the particle number, a further simplification ensues. In this case, each term in the Hamiltonian will consist of an equal number of annihilation and creation operators. Hence, the Hamiltonian is invariant under the transformation $a_k \rightarrow a_k e^{i\phi}$ and $a_k^\dagger \rightarrow a_k^\dagger e^{-i\phi}$ and we find that expectation values vanish unless

they concern operators with the same number of annihilation and creation operators so that X can be assumed to take the form

$$X = \begin{pmatrix} T & 0 & 0 & 0 & 0 \\ 0 & S & 0 & 0 & 0 \\ 0 & 0 & M & 0 & 0 \\ 0 & 0 & 0 & R & 0 \\ 0 & 0 & 0 & 0 & Q \end{pmatrix}. \quad (14)$$

Note that the elements of the matrices T, S are coefficients of the one-electron reduced density matrix (1-RDM), while the matrices M, R, Q are representations of the two-electron reduced density matrix (2-RDM) [9]. Restricting matrices of the form eq. (14) to be positive semi-definite is equivalent with the positivity of the individual matrices T, S and M, R, Q . This constraint on the matrices containing the second moments defines the 1-positivity conditions, while the positivity of the matrices containing the fourth moments generates the 2-positivity conditions [8, 9]. For concreteness we list the matrices comprising the constraints on a system composed of two spinless fermions and the additional assumption of particle number conservation in appendix B.

D. Exploiting Translational Invariance

So far we considered optimization problems for up to 2-positivity conditions which reduced the number of free variables to scale as N^4 . Apart from superselection rules and the particle number conservation, no specific assumption on the structure of the underlying Hamiltonian has been exploited such that in principle this procedure is applicable to a wide range of different systems, e.g. quantum chemistry calculations [13] or higher spacial dimensions [18]. Since our intention is to apply this method to large lattice models, we now discuss a further simplification of the system to decrease the number of parameters, namely the translational invariance. Translation invariance with periodic boundary conditions of the physical system reduces the number of free variables considerably. In fact for the second moments, i.e. for the matrices T, S and U , this can be exploited immediately as these matrices then obey $T_{i,j} = t_{i-j}$, etc. Consequently, for translational invariant systems with periodic boundary conditions, the matrices comprising the second moments become circulant matrices [19] and can therefore be diagonalized by a discrete Fourier transform

$$V_{k,l} = \frac{1}{\sqrt{N}} e^{-2\pi i(k-1)(l-1)/N} \quad (15)$$

for $k, l = 1, \dots, N$. Reducing the number of variables for the fourth moments is not so transparent as for the second moments, but manageable. To do so, note that the ordering of the entries of the matrices M, G, H, R, I and Q is not fixed but can be chosen arbitrarily. Remember that the indices of these matrices are of the form $[(k, l), (m, n)]$ for $k, l, m, n = 1, \dots, N$. Rearranging the two-digit indices labeling the rows and columns in the following way:

$$\begin{aligned} (k, l) = & \{ (1, 1), (2, 2), \dots, (N, N), \\ & (1, 2), (2, 3), \dots, (N, 1), \\ & (1, 3), (2, 4), \dots, (N, 2), \\ & \dots \\ & (1, N), (2, 1), \dots, (N, N-1) \}, \end{aligned} \quad (16)$$

the matrices comprising the fourth moments decompose in a block structure where each block can be diagonalized by the unitary matrix $V_{k,l}$ given above. For example, we find for the matrix M

$$M = U D_M U^\dagger, \quad (17)$$

where $U = \text{diag}(V, V, \dots, V)$ with V given by eq. (15) and

$$D_M = \begin{pmatrix} D^{1,1} & D^{1,2} & \dots & D^{1,N} \\ \vdots & \vdots & \ddots & \vdots \\ D^{N,1} & D^{N,2} & \dots & D^{N,N} \end{pmatrix} \quad (18)$$

with diagonal matrices $D^{i,j}$, $i, j = 1, \dots, N$. Hence, for translational invariant systems the matrices contained in set \mathcal{C} will have the following form:

$$X = \begin{pmatrix} D_T & D_{U^\dagger} & 0 & 0 & 0 \\ D_U & D_S & 0 & 0 & 0 \\ 0 & 0 & D_M & D_{G^\dagger} & D_{H^\dagger} \\ 0 & 0 & D_G & D_R & D_{I^\dagger} \\ 0 & 0 & D_H & D_I & D_Q \end{pmatrix}, \quad (19)$$

where the matrices denoting the second moments in the upper left block are diagonal matrices and the matrices considering the fourth moments have a block diagonal structure as described above. These considerations reduce the complexity of the constraints considerably and hence make the numerical study feasible for large systems.

E. Varying System Size for Higher Moments

As the treatment of the fourth-order correlations is computationally costly, it may in some cases be advantageous to consider fourth-order constraints on a smaller lattice, while the computationally cheaper second-order constraints are treated on the full lattice. In particular, while T, U and S are $N \times N$ matrices, M, G, H, R, I and Q are $N^2 \times N^2$ matrices leading to the unfavourable scaling. To overcome this issue, it might therefore be beneficial to restrict higher moments to smaller subsystems; that is, for the second moments one would consider system size N_2 , while for fourth moments one would consider N_4 where $N_2 \gg N_4$. Reducing the support of the matrices M, G, H, R, I and Q , on the one hand, will lead to disproportionate savings in memory and computation time and, on the other hand, will certainly lead to a decrease of the lower bounds.

III. THE PROJECTED GRADIENT ALGORITHM

So far we have described how to relax the original optimization problem eq. (1) and discussed how to include symmetries to reduce the number of free variables. Now we are going to describe an iterative algorithm for addressing this minimization problem. Recall the minimization problem we have to face, that is,

$$\begin{aligned} & \text{minimize} && E(X) \\ & \text{subject to} && X \in \mathcal{C} \end{aligned} \quad (20)$$

where again the function $E(X) : \mathcal{H} \rightarrow \mathbb{R}$ expresses $\langle H \rangle$ and the set \mathcal{C} consists of all positive semi-definite Hermitian matrices X fulfilling the constraints dictated by the (anti)-commutator relations and potential other symmetries of the system. Note that the function $E(X)$ is indeed linear in the elements of matrix X and hence could be written as

$$E(X) = \text{tr}[G \cdot X] + c, \quad (21)$$

where G is a Hermitian matrix and $c \in \mathbb{R}$. Furthermore, the set \mathcal{C} is the intersection of the convex cone consisting of positive semi-definite matrices and the affine set defined by the linear constraints on the entries. Hence \mathcal{C} is convex, too [11]. It is well known that this optimization problem is in fact a semi-definite programme and as such can be solved efficiently. Standard primal-dual interior point methods for semi-definite programming come with a considerable overhead in memory and time which makes their application to large systems challenging. Recently, two algorithms addressing the problem of ground state energy estimation were developed which scale better than the conventional methods [13, 14]. Here, we continue the research for efficient algorithms and formulate an alternative numerical method whose efficacy we demonstrate in later sections.

To address the issues described above, we suggest to solve the minimization problem by a projected gradient algorithm [20]. Since this algorithm relies heavily on projections onto convex sets, we need to define the latter. Let \mathcal{H} be a vector space and $\mathcal{A} \in \mathcal{H}$ be a convex set. The projection of $X \in \mathcal{H}$ onto \mathcal{A} is defined by

$$P_{\mathcal{A}}(X) = \text{argmin} \{ \|X - Y\|_F \mid Y \in \mathcal{A} \}, \quad (22)$$

where $\|\cdot\|_F : \mathcal{H} \rightarrow \mathbb{R}$ denotes the Frobenius or Hilbert-Schmidt norm. As the name suggests, the projection determines the nearest point in the convex set \mathcal{A} with respect to a given norm. With this, we claim that the solution of problem eq. (20) could be found by iterating

$$X_{k+1} = P_{\mathcal{C}}(X_k - \alpha G), \quad (23)$$

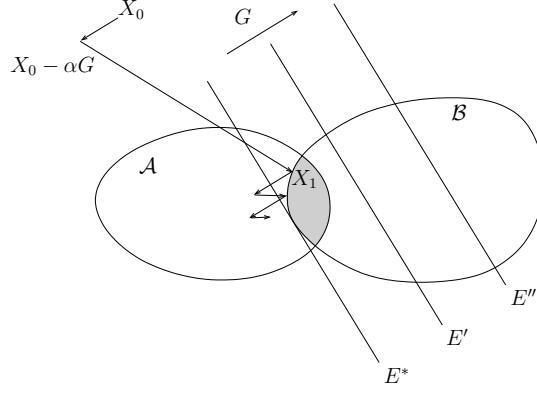


FIG. 1. Illustration of the gradient projection algorithm (23). The convex sets \mathcal{A} and \mathcal{B} have a nonempty intersection, which is shown shaded. The straight parallel lines clarify the level sets of the function $E(X)$ such that $E^* < E' < E''$, where E^* is the minimum value the function attains on the intersection. The projections are represented by arrows to demonstrate their directions. Note that these projections are computed using Dykstra's algorithm, which itself uses the projections onto \mathcal{A} and \mathcal{B} . The algorithm is initialized with the matrix X_0 .

where the algorithm is initialized by a Hermitian matrix $X_0 \in \mathcal{H}$ and $\alpha > 0$ is an arbitrary real and positive number. Furthermore, $P_{\mathcal{C}} : \mathcal{H} \rightarrow \mathcal{C}$ denotes the projection onto set \mathcal{C} , i.e. the intersection of the positive semi-definite Hermitian matrices and the set given by the linear constraints. The scheme of this algorithm is visualized in figure 1.

We find that the following lemma holds.

Lemma 1 *Let $E : \mathcal{H} \rightarrow \mathbb{R}$ be a linear function of the form $E(X) = \langle G, X \rangle + c$ where $G \in \mathcal{H}^{n \times n}$ and $c \in \mathbb{R}$ is a constant. Assume that the function $E(X)$ is lower bounded on the closed convex set \mathcal{C} , i.e. there is an $\bar{X} \in \mathcal{C}$ such that $E(\bar{X}) \leq E(X)$ for all $X \in \mathcal{C}$. Let $E_k = E(X_k)$ be the sequence determined by the iterations of the algorithm eq. (23). Then $E_k = E(X_k) \rightarrow E^*$ where $E^* \leq E(X)$ for all $X \in \mathcal{C}$.*

A detailed proof of this lemma is presented in appendix A.

So far we have shifted the problem of minimizing the function $E(X)$ with respect to the convex set \mathcal{C} to the problem of computing the projection onto \mathcal{C} . As pointed out above, this set is the intersection of two convex sets. This complicates the computation of the projection because in principle one has to solve the optimization problem eq. (22) in each step of the algorithm. The most obvious method at hand to overcome this difficulty is to use Dykstra's algorithm [17, 21]. Dykstra's algorithm is an elaborate technique to compute the projection of X onto the intersection of several convex sets by means of modified iterative projections onto the individual sets. To apply this procedure to the computation onto \mathcal{C} we need to determine the projections onto the set of positive semi-definite Hermitian matrices and onto the affine set given by the constraints imposed by the canonical (anti)-commutation relations and other symmetries. The former can be implemented straightforwardly. For the latter we were able to determine formulae for the submatrices in eq. (13) and thus gain a method to compute $P_{\mathcal{C}} : \mathcal{H} \rightarrow \mathcal{C}$ efficiently. Note that since \mathcal{C} is convex, $P_{\mathcal{C}}(X)$ is unique for all $X \in \mathcal{H}$ [17]. Further, the proposed algorithm finds a lower bound on the ground state energy by only considering boundary points of the feasible set and hence is slightly related to the algorithm proposed in [14].

There are two parameters defining the accuracy of the estimated lower bound. In principle the algorithm (23) converges for perfect projections $P_{\mathcal{C}}(X)$ to a matrix $X \in \mathcal{C}$ minimizing the function $E(X)$ in the limit $k \rightarrow \infty$, where k is the iteration number. As for all iterative algorithms we need to introduce stopping criteria, on the one hand, for the projected gradient algorithm and, on the other hand, for Dykstra's algorithm. For the latter we use the robust stopping criterion introduced in [21] and determined by a small number $\tau_{Dykstra} \in \mathbb{R}$. For the former one can exploit the fact that the dual problem of the optimization problem eq. (20), which we identify as the primal problem, can be solved using the projected gradient algorithm, too. Note that the primal problem can be written as

$$\begin{aligned} & \text{minimize} \quad \vec{g}^\dagger \vec{x} \\ & \text{subject to} \quad F_0 + \sum_i x_i F_i \geq 0, \end{aligned} \tag{24}$$

where we identify the matrices G and X with the vectors \vec{g} and \vec{x} such that $E(X) - c = \langle G, X \rangle = \vec{g}^\dagger \vec{x}$. The Hermitian matrices $F_0, \{F_i\}_i$ comprise the linear constraints dictated by the fermionic anti-commutator relations such that the matrix $F_0 + \sum_i x_i F_i$ is an element of the affine set for all $x_i \in \mathbb{R}$. The dual problem of eq. (24) is then for the dual

variable $Z \in \mathcal{H}$ [10]

$$\begin{aligned} & \text{maximize} && -\langle F_0, Z \rangle \\ & \text{subject to} && g_i = \langle F_i, Z \rangle \\ & && Z \geq 0. \end{aligned} \tag{25}$$

Note that the constraints of the dual problem force the solution Z^* to lie both in the positive semi-definite cone and in the affine set given by the linear equations. The intersection of these two convex sets is again convex and the dual problem becomes

$$\begin{aligned} & \text{minimize} && \langle F_0, Z \rangle \\ & \text{subject to} && Z \in \mathcal{D}, \end{aligned} \tag{26}$$

where \mathcal{D} denotes the intersection of the positive semi-definite cone with the set given by the linear constraints $g_i = \langle F_i, Z \rangle$. Finally, lemma 1 shows that the projected gradient algorithm can be applied to solve the dual problem as well.

The benefit of the formulation as a primal-dual problem is that the solution of the dual problem is a lower bound on the solution of the primal problem. Hence, solving the primal and dual problems simultaneously using the projected gradient algorithm provides us with an ingenious certificate for the accuracy of the solution. Consequently, one way to introduce a stopping criterion for the projected gradient algorithm is to initiate a small quantity $\tau_{\text{primal-dual}}$ and stop the computations when

$$E(X_k) - H(Z_k) \leq \tau_{\text{primal-dual}}, \tag{27}$$

where $H(Z) = -\langle F_0, Z \rangle + c$ is basically the dual function up to a constant. Note that eq. (27) is equivalent to the duality gap of the primal and dual problem and that for strictly feasible primal problems, i.e. for optimization problems of the form eq. (24) for which there exists an x such that $F_0 + \sum_i x_i F_i > 0$, the duality gap is zero if evaluated at optimal points [10]. For the optimization problem eq. (20), one can easily find a point in the feasible set which is strictly feasible. Consequently, we have $E(X^*) - H(Z^*) = 0$, where X^* and Z^* are solutions of the primal and dual problems, respectively, and hence eq. (27) suits perfectly as a stopping criterion. Due to the numerical cost to solve the primal and dual problems simultaneously, we decide on using a different stopping criterion for the projected gradient algorithm. We stop when the improvement per iteration step, i.e.

$$E(X_{k+1}) - E(X_k) \leq \tau_E, \tag{28}$$

falls below a predetermined small number $\tau_E \in \mathbb{R}$. We implemented both routines for systems considering only second moments and numerical experiments suggest that for a sufficiently small number τ_E the latter stopping criterion yields results that are as good as in the primal-dual formalism.

Another degree of freedom in the formulation of the algorithm is the step length α . Numerical examples propose that for the first step in the direction of the negative gradient α should be chosen to be large. For the remaining steps the computational cost could be reduced by choosing a small but appropriate value of α [22].

IV. APPLICATIONS

In the remainder of this work, we provide a number of numerical examples that demonstrate that the proposed method can be applied to rather large systems and is capable of extracting accurate information concerning the ground state energy as well as the correlation functions. The purpose of what follows is not to improve lower bounds on unknown ground state energies, but rather to analyse the behaviour of the proposed algorithm for finite but large lattice models.

A. Second Moments and the Ising Model

We begin our exposition of the numerical application of the algorithm with a simple, exactly solvable model for which the relaxation of the energy minimization problem has been analysed in great detail exploiting different numerical methods in [26]. The Hamiltonian we are considering as a first application of the projected gradient algorithm is given

by

$$\begin{aligned}
H = & j_c \sum_{k=1}^{N-1} \left(a_k^\dagger a_{k+1} + a_{k+1}^\dagger a_k \right) - j_c \left(a_N^\dagger a_1 + a_1^\dagger a_N \right) \\
& + j_c \sum_{k=1}^{N-1} \left(a_k^\dagger a_{k+1}^\dagger + a_{k+1} a_k \right) - j_c \left(a_N^\dagger a_1^\dagger + a_1 a_N \right) \\
& - hN + 2h \sum_{k=1}^N a_k^\dagger a_k,
\end{aligned} \tag{29}$$

where the operators describe spinless fermions. The ground state of this model coincides with that of the Ising model in a transverse magnetic field with periodic boundary conditions

$$H = j_c \sum_{k=1}^N \sigma_k^x \sigma_{k+1}^x - h \sum_{k=1}^N \sigma_k^z \tag{30}$$

when $j_c < 0$ and N is even. This can be seen from the connection of eq. (30) with eq. (29) via a Jordan-Wigner transformation [23, 24]. To analyse a translational invariant system with periodic boundary conditions, we simply change the signs of the second and the fourth terms in the Hamiltonian eq. (29).

Now we require only constraints that are due to second moments, i.e. we require that

$$X = \begin{pmatrix} T & U^\dagger \\ U & S \end{pmatrix} \geq 0 \tag{31}$$

and that the entries fulfil the constraints dictated by the anti-commutator relations. Solving the relaxed minimization problem eq. (20) by applying the projected gradient algorithm and comparison of the so-obtained lower bound to the exact ground state energy [23, 24]

$$E_0 = - \sum_{k=0}^{N-1} \left[\left(j_c \cos \left(\frac{\pi l_k}{N} \right) + h \right)^2 + \left(j_c \sin \left(\frac{\pi l_k}{N} \right) \right)^2 \right]^{1/2} \tag{32}$$

yields the results presented in table I. Note that the ground state energy for the system given by the Hamiltonian eq. (29) is computed by the formula (32) with parameter $l_k = 2k + 1$, while one chooses $l_k = 2k$ for the energy of a translational invariant system with periodic boundary conditions.

N	1000	5000	10000	50000	100000	500000
ΔE	$2.4 \cdot 10^{-8}$	$2.5 \cdot 10^{-8}$	$2.6 \cdot 10^{-8}$	$2.4 \cdot 10^{-8}$	$2.4 \cdot 10^{-8}$	$2.9 \cdot 10^{-8}$

TABLE I. Lower bounds on the ground state energies for the translational invariant Jordan-Wigner transformation of the Ising Hamiltonian with parameters $j_c = -1$ and $h = 1/2$ for N modes. The Hamiltonian is identical to eq. (29) but the signs of the second and fourth terms are changed to make the Hamiltonian translational invariant. The relative deviation ΔE is of the order of the parameters of the stopping criteria $\tau_{Dykstra} = \tau_E = 10^{-8}$, which limits the accuracy of this method. We initialized the algorithm with $\alpha = N$ for the first step and $\alpha \ll N$ for the remaining steps until convergence.

In fact, it can be shown directly that the ground state energy of the Ising model is *exactly* reproduced with the present method as the Ising model can be diagonalized by an orthogonal mode transformation mapping second moments to second moments, this making the transformed optimization problem straightforward to solve [25]. Thus the so obtained lower bounds are, in this case, superior to other approaches such as Anderson lower bounds. In particular, solving the optimization problem for the non-translational invariant Hamiltonian eq. (29) for even N and negative j_c yields the exact ground state energies for the translational invariant Ising Hamiltonian eq. (30). Even for this Hamiltonian one is able to solve the optimization problem for large particle numbers since one is only considering the 1-positivity conditions which scale linearly in the system size.

B. Fourth Moments and the Heisenberg model

While the Ising model is quasi-free it is of considerable interest to study interacting models, such as the Heisenberg Hamiltonian, that contain, for example, fourth-order fermionic operators (see [27] for the study of another interacting

fermionic system where a different numerical approach is employed). In the following, we consider the performance of the proposed algorithm in this situation. The Heisenberg Hamiltonian

$$H = J \sum_{i=1}^{N-1} [\sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y + \sigma_i^z \sigma_{i+1}^z] \quad (33)$$

can be transformed to spinless fermions by the Jordan-Wigner transformation. In this picture, the Hamiltonian reads

$$\begin{aligned} H = 2J \sum_{i=1}^{N-1} [a_{i+1}^\dagger a_i + a_i^\dagger a_{i+1}] - 2J \sum_{i=1}^{N-1} [a_{i+1}^\dagger a_{i+1} + a_i^\dagger a_i] \\ + 4J \sum_{i=1}^{N-1} a_i^\dagger a_{i+1}^\dagger a_{i+1} a_i + J \cdot (N-1). \end{aligned} \quad (34)$$

Now we are in a position to compare the exact ground state energies of the Heisenberg model with the lower bounds obtained by applying the suggested algorithm. Indeed, for the Heisenberg Hamiltonian with parameter $J = 1/2$ the lower bounds on the ground state energy lie very close to the actual energies. Table II presents some numerical results.

N	E_{exact}	Lower bound	Ratio
6	-4.9872	-4.9974	0.9979
8	-6.7499	-6.7739	0.9964
10	-8.5161	-8.5557	0.9954
12	-10.2842	-10.3412	0.9945
14	-12.0534	-12.1291	0.9938
16	-13.8235	-13.9189	0.9931
18	-15.5940	-15.7099	0.9926
20	-17.3649	-17.5021	0.9922

TABLE II. Lower bounds for the ground state energy of the Heisenberg Hamiltonian eq. (34) of N modes and with parameter $J = 1/2$. The exact ground state energy is obtained by diagonalizing the Hamiltonian in the spin picture. The parameters defining the stopping criteria for Dykstra's algorithm and for the projected gradient algorithm are $\tau_{Dykstra} = 10^{-4}$ and $\tau_E = 10^{-4}$. Increasing the precision of the algorithm by decreasing these parameters is accompanied by an improvement of the lower bounds and an increase of the computing time.

The range of applicability of the projected gradient algorithm can be extended considerably by analysing a system that is translationally invariant. To this end, we make the Hamiltonian in eq. (34) translationally invariant with periodic boundary conditions. The results for the ground state energy of this Hamiltonian are presented in the left-hand side of figure 2. As a reference, we depict the exact ground state energies for this model for up to $N = 22$. Additionally, the right-hand side of figure 2 presents the nearest-neighbour and next-nearest-neighbour correlations determined by the gradient projection algorithm.

To illustrate the rate of convergence of the algorithm for fourth moments, we present the improvement per iteration of the function $E(X_k) = \langle G, X_k \rangle + c$ for the translational invariant Jordan-Wigner transformed Heisenberg Hamiltonian describing $N = 50$ sites in figure 3. It is noticeable that the algorithm converges to a good estimation in only a few steps and that the most time-consuming factor is given by the projection onto the set \mathcal{C} .

Note that in all calculations we initialized the algorithm with the zero matrix since we wanted to test the algorithm without imposing prior knowledge. Generally, for translational invariant systems it is advantageous to complete the circulant submatrices obtained from a run with site number N_1 to circulant matrices describing N_2 sites and use these matrices as an initial matrix for $N_2 > N_1$.

Furthermore, since the matrices comprising the fourth moments in the translational invariant case decompose into blocks of diagonal matrices, it is possible to parallelize the projection onto the positive semi-definite matrices. In fact, this step in the computation is the most expensive since one has to diagonalize the matrices in each step. Parallelizing this computation might speed up the calculations. So far, no effort has been made to do so since our intention is rather to demonstrate that this technique yields good results.

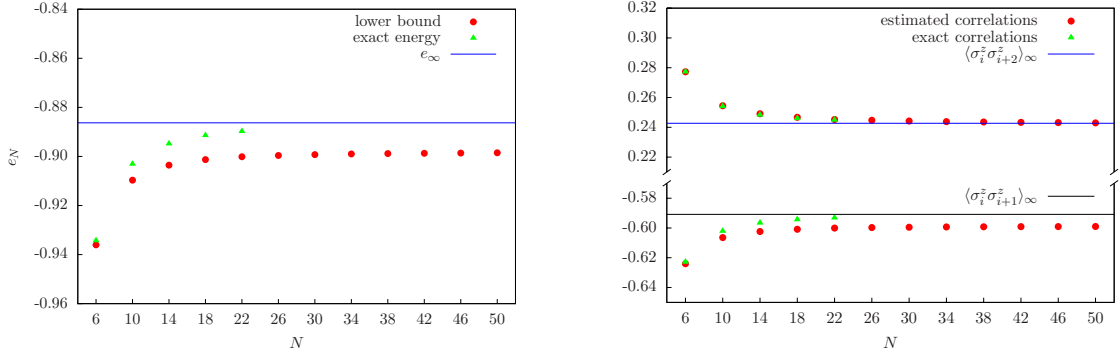


FIG. 2. The system under consideration is the Hamiltonian eq. (34) extended to a translational invariant system with periodic boundary conditions and parameter $J = 1/2$. The stopping criteria are initialized with $\tau_{Dykstra} = 10^{-4}$ and $\tau_E = 10^{-4}$. *Left:* lower bounds on the ground state energies $e_N = E_N/N$ per site. Circles indicate the estimated lower bounds obtained via the projected gradient algorithm. The exact energy per particle is illustrated as triangles for up to $N = 22$ particles. Additionally, the energy per particle in the thermodynamical limit $e_\infty = -4|J|(\ln 2 - 0.25)$ [23] is shown as a horizontal line to indicate where the exact energies converge to. *Right:* nearest- and next-nearest-neighbour correlations. Shown are the estimated correlations as circles, the exact correlations as triangles and the correlations in the thermodynamical limit as lines, i.e. $\langle \sigma_i^z \sigma_{i+1}^z \rangle_\infty = (1 - 4 \ln 2)/3 = -0.59086292$ and $\langle \sigma_i^z \sigma_{i+2}^z \rangle_\infty = (1 - 16 \ln 2 + 9\zeta(3))/3 = 0.24271908$ where $\zeta(3) = \frac{5}{2} \sum_{k=1}^{\infty} (-1)^{k-1} / [k^3 \binom{2k}{k}] \cong 1.2020569\dots$. The nearest-neighbour correlations are obtained by averaging over $\langle \sigma_i^\alpha \sigma_{i+1}^\alpha \rangle$ for $\alpha = x, y, z$, which is possible since the system is rotationally invariant. For the next-nearest-neighbour correlations, we proceeded the same way.

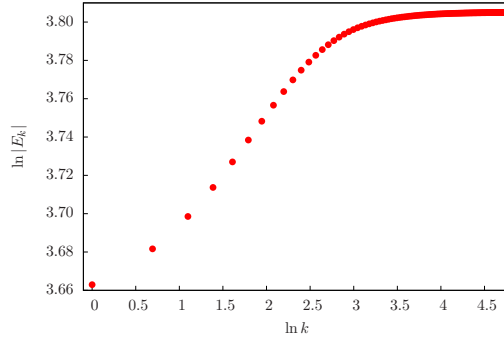


FIG. 3. Energies E_k for each iteration step for the Hamiltonian eq. (34) extended to a translational invariant system for $N = 50$ particles illustrating the good behaviour of the gradient projection algorithm with respect to the iteration step k . Note that the data is plotted double logarithmically.

V. CONCLUSIONS

In this work, we have presented a method for the determination of lower bounds on the ground state energy of quantum many-body systems. To this end, we followed a large body of work to relax the general ground state energy finding problem to obtain a semi-definite programme that involves a number of variables that scale polynomially with the system size. To overcome unfavourable scaling in time and memory of standard primal-dual interior point solvers, we combined a projected gradient algorithm with Dykstra's algorithm to obtain access to larger system sizes and proved its convergence. For quasi-free systems the method yields the exact ground state energies, while for general interacting models it provides very good lower bounds and direct access to correlation functions between sites. Different properties of the underlying system such as higher spatial dimensions [18], bosonic degrees of freedom [28], symmetries [29] and translational invariance can be included straightforwardly and the relative weight of higher order correlation functions can be adjusted to optimize for computational performance.

We hope that further investigations will help to optimize this approach and thus elucidate its power and limitations.

VI. ACKNOWLEDGEMENTS

We acknowledge support from the Bundesministerium für Bildung und Forschung (FK 01BQ1012), the EU Integrating Project Q-ESSENCE and the Alexander von Humboldt Foundation. Interesting discussions with K. Audenaert, J. Cai, M. Cramer, D.A. Mazziotti and M. Navascues are acknowledged.

Appendix A: Proof of Convergence

The analysis of lemma 1 requires some basic properties of the projection operators defined in eq. (22). The following results can be found, for example, in [17], theorems 4.1 and 5.5.

Lemma 2 *Let $P_{\mathcal{C}} : \mathcal{H} \rightarrow \mathcal{C}$ be the projection onto the convex set \mathcal{C} .*

1. *For all $Z \in \mathcal{C}$ and all $X \in \mathcal{H}$ we have $\langle P_{\mathcal{C}}(X) - X, Z - P_{\mathcal{C}}(X) \rangle \geq 0$.*
2. *For all $X, Y \in \mathcal{H}$ we have $\langle P_{\mathcal{C}}(X) - P_{\mathcal{C}}(Y), X - Y \rangle \geq 0$.*

The following proof is motivated by [20].

Proof of lemma 1

Before we start with the main proof, we need to establish the relation

$$\langle G, X_k - X_{k+1} \rangle \geq \frac{\|X_{k+1} - X_k\|_F^2}{\alpha} \quad (\text{A1})$$

for $\alpha \geq 0$ and where the sequence $\{X_k\}_k$ is defined via

$$X_{k+1} = P_{\mathcal{C}}(X_k - \alpha G), \quad (\text{A2})$$

i.e. the projected gradient algorithm. To verify relation (A1), note that with the first inequality of lemma 2, we find

$$\begin{aligned} & \langle X_{k+1} - X_k + \alpha G, X_k - X_{k+1} \rangle \\ &= \langle P_{\mathcal{C}}(X_k - \alpha G) - X_k + \alpha G, X_k - P_{\mathcal{C}}(X_k - \alpha G) \rangle \\ &\geq 0 \end{aligned} \quad (\text{A3})$$

since $X_k \in \mathcal{C}$. Hence we have

$$\langle \alpha G, X_k - X_{k+1} \rangle \geq \|X_k - X_{k+1}\|_F^2 \quad (\text{A4})$$

and relation (A1) is established. Note that with this expression the following holds for all $\alpha > 0$:

$$\langle G, X_{k+1} - X_k \rangle \leq 0. \quad (\text{A5})$$

In a first step, we show that the sequence $E(X_k)$ is monotonically decreasing. We find that

$$\begin{aligned} & \langle G, X_{k+1} \rangle \leq \langle G, X_k \rangle, \\ & \Leftrightarrow \langle G, X_{k+1} \rangle + c \leq \langle G, X_k \rangle + c, \\ & \Leftrightarrow E(X_{k+1}) \leq E(X_k), \end{aligned} \quad (\text{A6})$$

where $E(X) = \langle G, X \rangle + c$, which holds for all $\alpha \geq 0$. Hence the sequence $E(X_k)$ is monotonically decreasing. Since we assume that the function $E(X)$ is lower bounded on \mathcal{C} , we then conclude that this sequence converges for all $\alpha \geq 0$, i.e. $E(X_k) \rightarrow E^*$ for $k \rightarrow \infty$. It remains to be shown that E^* is indeed the minimum of the function in \mathcal{C} .

In a second step, we show that

$$\lim_{k \rightarrow \infty} \|X_{k+1} - X_k\|_F = 0. \quad (\text{A7})$$

Since $E(X)$ is linear in X , we have

$$\begin{aligned} E(X_k) &= \langle G, X_k \rangle + c + \langle G, X_{k+1} \rangle - \langle G, X_{k+1} \rangle \\ &= \langle G, X_{k+1} \rangle + c + \langle G, X_k - X_{k+1} \rangle \\ &= E(X_{k+1}) + \langle G, X_k - X_{k+1} \rangle \end{aligned} \quad (\text{A8})$$

and hence

$$\lim_{k \rightarrow \infty} E(X_k) = \lim_{k \rightarrow \infty} E(X_{k+1}) + \lim_{k \rightarrow \infty} \langle G, X_k - X_{k+1} \rangle. \quad (\text{A9})$$

Since the sequence $E(X_k)$ converges, i.e. $E(X_k) \rightarrow E^*$, we find that

$$\lim_{k \rightarrow \infty} \langle G, X_k - X_{k+1} \rangle = 0. \quad (\text{A10})$$

Inequality (A1) then yields

$$\begin{aligned} \lim_{k \rightarrow \infty} \langle G, X_k - X_{k+1} \rangle &\geq \lim_{k \rightarrow \infty} \frac{\|X_{k+1} - X_k\|_F^2}{\alpha}, \\ \Leftrightarrow 0 &\geq \lim_{k \rightarrow \infty} \frac{\|X_{k+1} - X_k\|_F^2}{\alpha} \end{aligned} \quad (\text{A11})$$

and hence

$$\lim_{k \rightarrow \infty} \|X_{k+1} - X_k\|_F^2 = 0. \quad (\text{A12})$$

Thus we have

$$\lim_{k \rightarrow \infty} \|X_{k+1} - X_k\|_F = 0. \quad (\text{A13})$$

In a third step, we show that the elements X_k are bounded. Let $\bar{X} \in \mathcal{C}$ be such that $E(\bar{X}) \leq E(X)$ for all $X \in \mathcal{C}$. Then

$$\begin{aligned} \|X_k - \bar{X}\|_F^2 &= \|X_k - X_{k+1} + X_{k+1} - \bar{X}\|_F^2 \\ &= \|X_k - X_{k+1}\|_F^2 + \|X_{k+1} - \bar{X}\|_F^2 \\ &\quad + 2\langle X_k - X_{k+1}, X_{k+1} - \bar{X} \rangle \\ &= \|X_k - X_{k+1}\|_F^2 + \|X_{k+1} - \bar{X}\|_F^2 \\ &\quad + 2\langle X_k - \alpha G + \alpha G - X_{k+1}, X_{k+1} - \bar{X} \rangle \\ &= \|X_k - X_{k+1}\|_F^2 + \|X_{k+1} - \bar{X}\|_F^2 \\ &\quad + 2\langle X_k - \alpha G - P_{\mathcal{C}}(X_k - \alpha G), P_{\mathcal{C}}(X_k - \alpha G) - \bar{X} \rangle \\ &\quad + 2\langle \alpha G, P_{\mathcal{C}}(X_k - \alpha G) - \bar{X} \rangle \\ &= \|X_k - X_{k+1}\|_F^2 + \|X_{k+1} - \bar{X}\|_F^2 \\ &\quad + 2\langle P_{\mathcal{C}}(X_k - \alpha G) - X_k + \alpha G, \bar{X} - P_{\mathcal{C}}(X_k - \alpha G) \rangle \\ &\quad + 2\alpha [E(X_{k+1}) - E(\bar{X})], \end{aligned} \quad (\text{A14})$$

since $\langle G, P_{\mathcal{C}}(X_k - \alpha G) - \bar{X} \rangle = \langle G, X_{k+1} - \bar{X} \rangle = E(X_{k+1}) - E(\bar{X})$. Note that with lemma 2 and the fact that $E(\bar{X}) \leq E(X)$ for all $X \in \mathcal{C}$, one can show that the last two terms are greater than or equal to zero. Therefore

$$\|X_k - \bar{X}\|_F^2 \geq \|X_k - X_{k+1}\|_F^2 + \|X_{k+1} - \bar{X}\|_F^2 \quad (\text{A15})$$

and consequently

$$\|X_{k+1} - \bar{X}\|_F^2 \leq \|X_k - \bar{X}\|_F^2 - \|X_k - X_{k+1}\|_F^2, \quad (\text{A16})$$

which leads to

$$\|X_{k+1} - \bar{X}\|_F^2 \leq \|X_k - \bar{X}\|_F^2. \quad (\text{A17})$$

We conclude that the elements X_k are bounded. As a consequence we know that, since \mathcal{C} is closed, there is at least one converging subsequence $X_{n_k} \rightarrow X^*$ for $k \rightarrow \infty$, where X^* is an accumulation point of the sequence $\{X_k\}_k$.

In a last step, we show that for any accumulation point X^* it holds that $E(X^*) \leq E(X)$ for all $X \in \mathcal{C}$. With lemma 2, we find for all $Z \in \mathcal{C}$

$$\begin{aligned} \langle X_{k+1} - X_k + \alpha G, Z - X_{k+1} \rangle &\geq 0, \\ \Leftrightarrow \alpha \langle G, X_{k+1} - Z \rangle &\leq \langle X_{k+1} - X_k, Z - X_{k+1} \rangle. \end{aligned} \quad (\text{A18})$$

Further, we have

$$\begin{aligned}
& \alpha \langle G, X_{k+1} - Z \rangle \\
& \leq \langle X_{k+1} - X_k, Z - X_{k+1} \rangle \\
& \leq \langle X_{k+1} - X_k, Z - X_k + X_k - X_{k+1} \rangle \\
& \leq \langle X_{k+1} - X_k, Z - X_k \rangle + \langle X_{k+1} - X_k, X_k - X_{k+1} \rangle \\
& \leq \langle X_{k+1} - X_k, Z - X_k \rangle - \|X_{k+1} - X_k\|_F^2 \\
& \leq \langle X_{k+1} - X_k, Z - X_k \rangle \\
& \leq \|X_{k+1} - X_k\|_F \cdot \|Z - X_k\|_F,
\end{aligned} \tag{A19}$$

where we exploited the Cauchy-Schwarz inequality. Finally, we find that

$$\begin{aligned}
\langle G, X_k - Z \rangle &= \langle G, X_k - X_{k+1} \rangle + \langle G, X_{k+1} - Z \rangle \\
&\leq \langle G, X_k - X_{k+1} \rangle + \frac{1}{\alpha} \|X_{k+1} - X_k\|_F \cdot \|Z - X_k\|_F
\end{aligned} \tag{A20}$$

holds for all $Z \in \mathcal{C}$. Now choose a converging subsequence X_{n_k} such that $X_{n_k} \rightarrow X^*$. We already know that $\langle G, X_{n_k} - X_{n_k+1} \rangle \rightarrow 0$ due to equation (A10). Further, equation (A13) claims that $\|X_{n_k+1} - X_{n_k}\|_F \rightarrow 0$ for $k \rightarrow \infty$. Additionally, since the sequence $\{X_k\}_k$ is bounded, we know that for $\bar{X} \in \mathcal{C}$, where $E(\bar{X}) \leq E(X)$ for all $X \in \mathcal{C}$,

$$\begin{aligned}
\|Z - X_{n_k}\|_F &= \|Z - \bar{X} + \bar{X} - X_{n_k}\|_F \\
&\leq \|Z - \bar{X}\|_F + \|\bar{X} - X_{n_k}\|_F \\
&\leq \|Z - \bar{X}\|_F + \|\bar{X} - X_0\|_F \\
&\leq C(Z),
\end{aligned} \tag{A21}$$

where we exploited that the sequence $\|\bar{X} - X_{n_k}\|_F$ is upper bounded by $\|\bar{X} - X_0\|_F$, see equation (A17), and where $C(Z)$ is a constant which may depend on arbitrary $Z \in \mathcal{C}$. Hence we find for all $Z \in \mathcal{C}$ and all accumulation points X^*

$$\begin{aligned}
\lim_{k \rightarrow \infty} \langle G, X_{n_k} - Z \rangle &\leq 0, \\
&\Rightarrow \langle G, X^* - Z \rangle \leq 0, \\
&\Rightarrow \langle G, X^* \rangle \leq \langle G, Z \rangle, \\
&\Rightarrow \langle G, X^* \rangle + c \leq \langle G, Z \rangle + c, \\
&\Rightarrow E(X^*) \leq E(Z).
\end{aligned} \tag{A22}$$

Since we know that the sequence $E(X_k)$ converges, we can choose the subsequence defined by $X_{n_k} \rightarrow X^*$ to see that $E(X_{n_k}) \rightarrow E^* = E(X^*)$. With the argumentation above, we know that $E^* \leq E(Z)$ for all $Z \in \mathcal{C}$ and hence the algorithm (A2) finally ends up at a point minimizing the function $E(X)$ in the convex set \mathcal{C} . \square

Appendix B: Constraints for a Fermionic System for up to 2-positivity Conditions

Here we list the constraints on the matrices for the second and fourth moments which arise by applying the fermionic anti-commutator relations and further specify the matrices under the additional assumption of particle number conservation for a system composed of two spinless fermions. For the second moments of a general, fermionic system, we find

$$\begin{aligned}
T_{k,l} &= \langle a_k^\dagger a_l \rangle, \\
S_{k,l} &= \langle a_k a_l^\dagger \rangle = \delta_{k,l} - T_{l,k}, \\
U_{k,l} &= \langle a_k a_l \rangle = -U_{l,k}.
\end{aligned} \tag{B1}$$

The fourth moments yield

$$\begin{aligned}
M_{kl,mn} &= \langle a_k^\dagger a_l^\dagger a_n a_m \rangle = -M_{lk,mn} = -M_{kl,nm} = M_{lk,nm}, \\
G_{kl,mn} &= \langle a_k^\dagger a_l a_n a_m \rangle = -G_{kl,nm} = G_{km,nl} \\
&= -G_{kn,ml} = -G_{km,ln} = G_{kn,lm}, \\
H_{kl,mn} &= \langle a_k a_l a_n a_m \rangle \\
&= -H_{kl,nm} = H_{km,nl} = -H_{km,ln} = H_{kn,lm} \\
&= -H_{kn,ml} = -H_{nl,mk} = H_{nk,ml} = -H_{nm,kl} \\
&= H_{nl,km} = -H_{nk,lm} = H_{nm,lk} = -H_{lk,mn} \\
&= H_{lk,nm} = -H_{lm,nk} = H_{lm,kn} = -H_{ln,km} \\
&= H_{ln,mk} = -H_{ml,kn} = H_{ml,nk} = -H_{mk,nl} \\
&= H_{mk,ln} = -H_{mn,lk} = H_{mn,kl}, \\
R_{kl,mn} &= \langle a_k^\dagger a_l a_n^\dagger a_m \rangle = \delta_{l,n} T_{k,m} - M_{kn,ml}, \\
I_{kl,mn} &= \langle a_k a_l a_n^\dagger a_m \rangle = \delta_{l,n} U_{k,m} - \delta_{k,n} U_{l,m} + G_{nk,ml} \text{ and} \\
Q_{kl,mn} &= \langle a_k a_l a_n^\dagger a_m^\dagger \rangle \\
&= \delta_{l,n} \delta_{k,m} - \delta_{l,n} T_{m,k} - \delta_{k,n} \delta_{l,m} + \delta_{k,n} T_{m,l} \\
&\quad + \delta_{l,m} T_{n,k} - \delta_{k,m} T_{n,l} + M_{nm,lk}.
\end{aligned} \tag{B2}$$

Assuming particle number conservation for a system composed of two spinless fermions, the matrices containing the second moments read

$$T = \begin{pmatrix} \langle a_1^\dagger a_1 \rangle & \langle a_1^\dagger a_2 \rangle \\ \langle a_1^\dagger a_2 \rangle^* & \langle a_2^\dagger a_2 \rangle \end{pmatrix}, \quad S = \begin{pmatrix} 1 - \langle a_1^\dagger a_1 \rangle & -\langle a_1^\dagger a_2 \rangle^* \\ -\langle a_1^\dagger a_2 \rangle & 1 - \langle a_2^\dagger a_2 \rangle \end{pmatrix},$$

while the matrices for the fourth moments are

$$\begin{aligned}
M &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \langle a_1^\dagger a_2^\dagger a_2 a_1 \rangle & -\langle a_1^\dagger a_2^\dagger a_2 a_1 \rangle & 0 \\ 0 & -\langle a_1^\dagger a_2^\dagger a_2 a_1 \rangle & \langle a_1^\dagger a_2^\dagger a_2 a_1 \rangle & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\
R &= \begin{pmatrix} \langle a_1^\dagger a_1 \rangle & 0 & \langle a_1^\dagger a_2 \rangle & \langle a_1^\dagger a_2^\dagger a_2 a_1 \rangle \\ 0 & \langle a_1^\dagger a_1 \rangle - \langle a_1^\dagger a_2^\dagger a_2 a_1 \rangle & 0 & \langle a_1^\dagger a_2 \rangle \\ \langle a_1^\dagger a_2 \rangle^* & 0 & \langle a_2^\dagger a_2 \rangle - \langle a_1^\dagger a_2^\dagger a_2 a_1 \rangle & 0 \\ \langle a_1^\dagger a_2^\dagger a_2 a_1 \rangle & \langle a_1^\dagger a_2 \rangle^* & 0 & \langle a_2^\dagger a_2 \rangle \end{pmatrix}, \\
Q &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 + \langle a_1^\dagger a_2^\dagger a_2 a_1 \rangle - \langle a_1^\dagger a_1 \rangle - \langle a_2^\dagger a_2 \rangle & -1 - \langle a_1^\dagger a_2^\dagger a_2 a_1 \rangle + \langle a_1^\dagger a_1 \rangle + \langle a_2^\dagger a_2 \rangle & 0 \\ 0 & -1 - \langle a_1^\dagger a_2^\dagger a_2 a_1 \rangle + \langle a_1^\dagger a_1 \rangle + \langle a_2^\dagger a_2 \rangle & 1 + \langle a_1^\dagger a_2^\dagger a_2 a_1 \rangle - \langle a_1^\dagger a_1 \rangle - \langle a_2^\dagger a_2 \rangle & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.
\end{aligned} \tag{B3}$$

The presented matrices T , S , M , R and Q now satisfy the fermionic anti-commutator relations for two spinless fermions. For a system where the particle number is conserved the constraint on the relaxed optimization problem eq. (12) is that these matrices are all simultaneously positive semi-definite, which generates the 1- and 2-positivity conditions.

-
- [1] S.R. White, Phys. Rev. Lett. **69**, 2863 (1992).
[2] U. Schollwoeck, Ann. Phys. **326**, 96 (2011).
[3] S. Östlund and S. Rommer, Phys. Rev. Lett. **75**, 3537 (1995); S. Rommer and S. Östlund, Phys. Rev. B **55**, 2164 (1997).

- [4] P. W. Anderson, Phys. Rev. **83**, 1260 (1951).
- [5] K. Husimi, Proc. Phys. Math. Soc. Japan **22**, 264 (1940); P.O. Löwdin, Phys. Rev. **97**, 1474 (1955); J.E. Mayer, Phys. Rev. **100**, 1579 (1955).
- [6] A.J. Coleman, Rev. Mod. Phys. **35**, 668 (1963).
- [7] C. Garrod and J. Percus, J. Math. Phys. **5**, 1756 (1964).
- [8] D.A. Mazziotti and R.M. Erdahl, Phys. Rev. A **63**, 042113 (2001).
- [9] D.A. Mazziotti, Ed., *Reduced-density-matrix mechanics: with application to many-electron atoms and molecules*, Advances in Chemical Physics, **134**, Wiley, New York, 2007.
- [10] L. Vandenberghe and S. Boyd, SIAM Rev. **38**, 49-95 (1996).
- [11] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, 2009.
- [12] S. Pironio, M. Navascues and A. Acin, SIAM J. Optim. **20**, 2157 (2010); M. Navascues, S. Pironio and A. Acin, New J. Phys. **10**, 073013 (2008).
- [13] D.A. Mazziotti, Phys. Rev. Lett. **93**, 213001 (2004).
- [14] D.A. Mazziotti, Phys. Rev. Lett. **106**, 083001 (2011).
- [15] J. Dattorro, *Convex Optimization & Euclidean Distance Geometry*, MeBoo Publishing, Palo Alto, 2005.
- [16] T. Baumgratz, Diploma Thesis, University Ulm, Germany 2010.
- [17] F. Deutsch, *Best approximation in inner product spaces*, Springer, New York, 2001.
- [18] A.V. Sinitskiy, L. Greenman, D.A. Mazziotti, J. Chem. Phys. **133**, 014104 (2010).
- [19] R.A. Horn and C.R. Johnson, *Matrix analysis*, Cambridge University Press, Cambridge, 1991.
- [20] P.H. Calamai and J.J. Moré, Math. Prog. **39**, 93-116 (1987).
- [21] E.G. Birgin and M. Raydan, SIAM J. Sci. Comput. **26**, 1405-1414 (2005).
- [22] Indeed, it seems that $\lim_{\alpha \rightarrow \infty} P_{\mathcal{C}}(X - \alpha G) = X^*$ holds, where $X \in \mathcal{H}$ is the initial matrix and $X^* \in \mathcal{C}$ is such that it minimizes the function $E(X)$ on the convex set \mathcal{C} , i.e. $E(X^*) \leq E(X)$ for all $X \in \mathcal{C}$.
- [23] M. Takahashi, *Thermodynamics of one-dimensional solvable models*, Cambridge University Press, 1999.
- [24] M. Reuter, *Entanglement Properties of Strongly Correlated Spin Systems*, PhD Thesis, Imperial College London, 2008.
- [25] J. Cai, Private Communication (2011).
- [26] C.A. Schwerdtfeger and D.A. Mazziotti, J. Chem. Phys. **130**, 224102 (2009).
- [27] J.R. Hammond and D.A. Mazziotti, Phys. Rev. A **73**, 062505 (2006).
- [28] G. Gidofalvi and D.A. Mazziotti, Phys. Rev. A **69**, 042511 (2004).
- [29] G. Gidofalvi and D.A. Mazziotti, Phys. Rev. A **72**, 052505 (2005).